



INNOVATION

Get the Good Stuff!

Melanie Gardner, AgNIC Coordinator, NAL
Vern Chapman, AgNIC IT, NAL
USAIN Conference
Wooster, Ohio
April 30, 2008

Agenda

- What are the problems?
- What is AgNIC's answer?
 - What is OAI & is it effective?
 - Is this sustainable?
 - Lessons learned

What are the problems?

■ Problem #1:

■ Cross searching databases and “interoperability”

- Expensive and time-consuming to build filters for multiple databases
- Federated searching depends on one or more systems at the other end – impacting response time
- Independent systems do not necessarily use the same named fields – creating a mix of information and a potential point of confusion for the user

■ Problem #2:

- How to offer the public user/customer more full-text

Thinking “outside the box”

- **INSIDE THE BOX** - Consider the traditional options
 - Build metadata
 - Hyper-link to sites
 - Content Management Systems
 - Rely on Google, Yahoo, etc.
- **OUTSIDE THE BOX** - What else can we do?
 - Identify full-text
 - Reuse what works
 - Go get it!

AgNIC's Answer: OAI

- What is OAI?

- Open Archives Initiative – OAI – is a “protocol which is intended to supply and promote an application-independent interoperability framework for communities involved in publishing content on the Web”.
- Effort began in 2000 and launched in January of 2001

OAI attempts to address

- Interoperability and exchange issues
- Exposure/access to targeted, openly available full-text

Is OAI Effective?

■ Yes and no

- Yes because it works!
- No because it is not adequate
 - Many digital objects are broken into parts. OAI is not optimal for aggregated digital objects.

■ Options

- Other protocols such as Z39.50, MARC, AGRIS, etc. – none of which are optimal – some are ubiquitous for databases but not in repositories
- New effort called OAI-ORE, or Open Archives Initiative – Object Reuse Exchange
 - allows distributed repositories to exchange information about their constituent digital objects

How Has AgNIC Applied OAI?

- *For the pilot project, we:*
 - Targeted repositories
 - Targeted collections within repositories
 - Performed technical work
 - Develop and run scripts that check for URLs, descriptions, format, and titles – essential pieces of descriptions [metadata]
 - For example - to normalize formats such as html, pdf, etc.
 - Establish OAI routines to harvest periodically

Why?

- Able to target appropriate collections
- Ease of implementation
 - Many applications have OAI “built in”
- Automated harvesting – makes this *nearly* maintenance-free

“AgOAI”

- 14 repositories identified, dozens of collections' metadata harvested
- Nearly 30,000 metadata records harvested
- Updated on different schedules according to number of items in the collections
- Added “subject search” feature
- We need more repositories and collections!



[Home](#) | [About](#) | [Partners](#) | [Browse](#) | [News](#) | [Calendar](#) | [AgOAI](#) | [Toolkit](#) | [Members](#) | [Contact Us](#)

SOURCES

- [African Journals Online](#)
- [AgSpace](#)
- [Aquatic Commons](#)
- [Auburn University Digital Library](#)
- [Library of Congress American Memory](#)
- [MINDS @ UW](#)
- [Organic Eprints](#)
- [Programa CyberTesis. Universidad de Chile](#)
- [Publication of Archival, Library and Museum Materials \(PALMM\)](#)
- [Purdue e-Pubs](#)
- [ScholarSpace at University of Hawaii at Manoa](#)
- [The Digital Library for Earth System Education \(DLESE\)](#)
- [The TAMU Digital Repository](#)
- [Washington State University Research Exchange](#)

This section contains full-text, freely available materials from the sources listed on the left. This collection is updated on a regular basis.

[Credits](#) | [Disclaimer](#) | [Contact Us](#) | [Accessibility](#) | [Site Map](#) | [Log Out](#)

Copyright 2000 - 2008 by AgNIC

http://txspace.tamu.edu/

AgOAI — AgNIC Alliance Repository Home

TEXAS A&M LIBRARIES Digital

Library → Digital Library → Repository Home Login

Search Repository
[Advanced Search](#)

▼ Browse

All of the Repository

- Communities & Collections
- By Issue Date
- Authors
- Titles
- Subjects

► My Account

MODELING HIGH-GENUS SURFACES
 A Dissertation by Vinod Srinivasan

Online Theses and Dissertations [\[+\] About this image](#)

What is a repository?

The Texas A&M Repository is an digital service that collects, preserves, and distributes the scholarly output of the university. The repository is an important tool to facilitate scholarly communication and preserve the scholarly legacy of Texas A&M faculty. Based on the principles of Open Access, the Texas A&M Repository contains many types of content, from electronic theses and dissertations to digitized maps from the Library's holdings.

Communities in the Repository

Select a community to browse its collections.

- Colleges & Schools
- Programs, Centers, and Institutes
- Special Collections
- State Agencies
- Texas A&M University Libraries
- Texas A&M University Press Consortium

Giving to the Libraries Texas A&M University Employment Opportunities Contact Us

Start | Inbox - Microsoft ... | Questions about R... | RE: Questions abo... | Get the Good Stuff! | AgNIC Services Up... | Repository Hom... | 2:48 PM

http://www.agnic.org/agoai Iowa

Home About Partners Browse News Calendar AgOAI Toolkit Contact Us

Manatee Search

SOURCES

- African Journals Online
- AgSpace
- Aquatic Commons
- Auburn University Digital Library
- Library of Congress American Memory
- MINDS @ UW
- Organic Eprints
- Programa CyberTesis. Universidad de Chile
- Publication of Archival, Library and Museum Materials (PALMM)
- Purdue e-Pubs
- ScholarSpace at University of Hawaii at Manoa
- The Digital Library for Earth System Education (DLESE)
- The TAMU Digital Repository
- Washington State University Research Exchange

Record 1 of 6

Title	/ edited by Thomas J. O'Shea, Bruce B. Ackerman, and H. Franklin Percival.
Identifier	http://purl.fcla.edu/fcla/dl/UF00000121.jpg
Identifier	http://purl.fcla.edu/fcla/dl/UF00000121.pdf
Description	Proceedings of a workshop held at the Austin Cary Forest of the University of Florida, Feb. 4-6, 1992.
Subject	Manatees--Florida--Congresses. ; Rare mammals--Florida--Congresses. ; West Indian manatee--Florida--Congresses.
Contributor	O'Shea, Thomas J., Ackerman, Bruce B. (Bruce Bennet), Percival, H. Franklin (Henry Franklin), United States. National Biological Service., PALMM (Project), Technical Workshop on Manatee Population Biology (February 4-6, 1992 : Gainesville, Fla.)
Publisher	U.S. Dept. of the Interior, National Biological Service
Date]
Provider	Publication of Archival, Library and Museum Materials (PALMM)

Record 2 of 6

Title	Ground-water resources data of Charlotte, De Soto, and Hardee Counties, Florida / by M.I. Kaufman and N.P. Dion.
Creator	Kaufman, Matthew I.
Identifier	http://purl.fcla.edu/UF/lib/GroundwaterResourcesDataOfCharlotte.jpg
Identifier	http://purl.fcla.edu/UF/lib/GroundwaterResourcesDataOfCharlotte.pdf
Description	"Prepared by the United States Geological Survey in cooperation with the Division of Geology, Florida Board of Conservation and the Southwest Florida Water Management District."
Subject	Groundwater--Florida--Charlotte County. ; Groundwater--Florida--De Soto County. ; Groundwater--Florida--Hardee County. ; Charlotte--12015 ; De Soto--12027 ; Hardee--12049
Contributor	Dion, N. P., Geological Survey (U.S.), Southwest Florida Water Management District (Fla.), PALMM (Project)
Publisher	State of Florida, Division of Geology
Date	1968
Provider	Publication of Archival, Library and Museum Materials (PALMM)

Record 3 of 6

Title	/ Roland M. Harper.
Creator	Harper, Roland M. (Roland McMillan), 1878-1966.
Identifier	http://purl.fcla.edu/fcla/dl/UF00000212.jpg
Identifier	http://purl.fcla.edu/fcla/dl/UF00000212.pdf

Find: Manatee Match case

Is this sustainable?

- Yes and no.
 - This is a very easy way to provide access to open access full-text
 - We do not have to manage the files or the metadata!
 - Need help identifying additional repositories
 - Need help identifying repository collections
 - New collections seem to be coming online fast

Lessons Learned

- AgNIC seems to be in front of the curve -
 - Many of the repositories we began harvesting are now adding collections – we have to go back and look at the new collections
 - Need to have discussions about the quality of repository metadata, or maybe it is not important?
 - Everyone has their own rules for metadata
 - We made it work, but some have not been able to make OAI work on their systems
 - We should build a committee, community, or such to further build out such a service

More ...

- Use of a single, or *selected* controlled vocabularies would make for better topic organization
- Begin to answer some of the questions that have come up about:
 - Metadata quality
 - Format standardization
 - Character encoding – UTF-8

INNOVATION

Questions?
Comments?

Melanie Gardner mgardner@nal.usda.gov

Vernon Chapman vchapman@nal.usda.gov